

QUANTUM TOMOGRAPHY UNDER PRIOR INFORMATION

TEIKO HEINOSAARI[‡], LUCA MAZZARELLA[‡], AND MICHAEL M. WOLF[‡]

ABSTRACT. We provide a detailed analysis of the question: how many measurement settings or outcomes are needed in order to identify a quantum system which is constrained by prior information? We show that if the prior information restricts the system to a set of lower dimensionality, then topological obstructions can increase the required number of outcomes by a factor of two over the number of real parameters needed to characterize the system. Conversely, we show that almost every measurement becomes informationally complete with respect to the constrained set if the number of outcomes exceeds twice the Minkowski dimension of the set. We apply the obtained results to determine the minimal number of outcomes of measurements which are informationally complete with respect to states with rank constraints. In particular, we show that $4d - 4$ measurement outcomes (POVM elements) is enough in order to identify all pure states in a d -dimensional Hilbert space, and that the minimal number is at most $2\log_2(d)$ smaller than this upper bound.

1. INTRODUCTION

Quantum tomography aims at identifying quantum systems. In order to achieve this, we use measurement data often supplemented by prior information. In this work we consider cases where prior information effectively reduces the dimensionality, i.e., the number of parameters which are necessary to characterize the state of a system. Physically, one may think of scenarios of interferometry, process tomography or parameter estimation, where one prepares the initial state which then evolves depending on a certain number of unknown parameters before one measures the final system. Effective reductions of the number of parameters can also be due to a constraining symmetry or fixed energy or particle number.

We are interested in determining the minimal number of measurement outcomes or measurement settings which are required for identifying a state taken from such a reduced set. Clearly, if states in the considered set are parameterized by a certain number of independent real parameters, then we need at least this number of measurement outcomes or binary measurements in order to pinpoint the state. As an example take the manifold of pure states in a d -dimensional Hilbert space. Their description requires $2d - 2$ real parameters, as opposed to $d^2 - 1$ real parameters needed to describe an arbitrary density matrix. So if we want to determine a pure state by a single measurement with m outcomes, $m \sim 2d$ seems necessary and $m \sim d^2$ appears to be achievable. The question about the smallest possible m in this particular example has been addressed in a number of publications [1, 2, 3, 4, 5], but the answer has remained somewhat elusive, so far.

A related problem has been addressed based on *compressed sensing* ideas, where it has been shown [6] that for $d \times d$ density matrices of rank r , $m = O(dr \log(d)^2)$ binary measurements are sufficient in order to identify the state with high probability. In this light we emphasize that our focus lies on schemes which identify the system unambiguously and deterministically. In this work we disregard questions regarding the robustness of the schemes, the complexity

of the necessary post-processing of the measurement data or the ability to verify the assumed prior information.

Outline. Typically, when we want to know the minimal number of measurement outcomes related to some subset of states, we need to independently consider upper and lower bounds. In Section 3 we concentrate on upper bounds, while in Section 4 we provide some methods to obtain lower bounds. The upper bounds are of geometric or algebraic nature and often come with concrete constructions. The lower bounds are mainly topological in nature. They are based on the observation that any measurement which is informationally complete when supplemented by prior information is a mapping into the space of measurement outcomes which preserves topological invariants. In the case of pure states the upper and lower bounds essentially match which enables us for instance to show that for a minimal informational complete measurement $m \sim 4d$ up to an additive logarithmic correction.

Notation. In this paper \mathcal{H} is a fixed *finite* dimensional complex Hilbert space. We denote by $\mathcal{L}(\mathcal{H})$ the set of linear operators on \mathcal{H} and we will write $\|L\|$ for the operator norm which coincides with the largest singular value of $L \in \mathcal{L}(\mathcal{H})$. A positive operator $\varrho \in \mathcal{L}(\mathcal{H})$ having trace one is a *density operator*, also referred to as *state*, and we denote by $\mathcal{S}(\mathcal{H})$ the set of all states on \mathcal{H} .

2. QUANTUM TOMOGRAPHY PREREQUISITES

In this section we will introduce and analyze the linear algebra framework for quantum tomography. Later, in Section 4, we will then add a topological perspective.

The goal of quantum tomography is to identify an unknown quantum state from the statistics of measurement outcomes. We will consider tomographic schemes where a fixed (as opposed to an adaptive) measurement setting is chosen, and we will refrain from considering errors for instance caused by finite statistics. In particular, in the following "measurement" always refers to a full statistical experiment rather than to a single-shot experiment.

Our analysis deals with several different albeit closely related approaches: 1) single measurements with many outcomes; 2) several measurements with possibly fewer outcomes; 3) several measurements where only expectation values are considered.

Quantum measurements are generally described by *positive operator valued measures* (POVMs) [7, 8]. Since we are eventually interested in the minimal required number of outcomes, we will restrict to POVMs with finitely many outcomes. For $n \in \mathbb{N}$ outcomes such a POVM is characterized by a collection of positive operators $\{A_j\}_{j=1}^n := \mathbf{A} \subset \mathcal{L}(\mathcal{H})$ satisfying $\sum_{j=1}^n A_j = \mathbb{1}$. Notice that a POVM with n outcomes is already determined by $n - 1$ operators; the last operator A_n is given by $A_n = \mathbb{1} - \sum_{j=1}^{n-1} A_j$.

For a POVM \mathbf{A} and a state (density matrix) ϱ , we denote by $\varrho^{\mathbf{A}}$ the corresponding probability distribution of measurement outcomes. It is given by the formula

$$\varrho^{\mathbf{A}}(j) = \text{tr}[\varrho A_j], \quad j = 1, \dots, n.$$

Again, due to normalization the last component is determined by the others via $\varrho^{\mathbf{A}}(n) = 1 - \sum_{j=1}^{n-1} \varrho^{\mathbf{A}}(j)$. A POVM with n outcomes therefore induces a continuous, linear map from $\mathcal{S}(\mathcal{H})$ into \mathbb{R}^n (or \mathbb{R}^{n-1} if we disregard the somewhat irrelevant last component). In the most favorable case $\varrho_1^{\mathbf{A}} \neq \varrho_2^{\mathbf{A}}$ for all pairs of different states ϱ_1 and ϱ_2 ; in this case \mathbf{A} is called *informationally complete* [9], [10] and the respective map $\varrho \mapsto \varrho^{\mathbf{A}}$ is injective on $\mathcal{S}(\mathcal{H})$. In this work we are interested in cases where prior information, or premise, limits the possible

initial states of the system. In other words, we want to identify ϱ not necessarily among all states but within some subset $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ of states. This motivates the following definition.

Definition 1 (\mathcal{P} -informational completeness). Let $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ be a subset of density matrices. A POVM \mathbf{A} is called *informationally complete w.r.t. \mathcal{P}* if $\varrho \mapsto \varrho^{\mathbf{A}}$ is injective on \mathcal{P} , i.e.,

$$\forall \varrho_1, \varrho_2 \in \mathcal{P} : \quad \varrho_1^{\mathbf{A}} = \varrho_2^{\mathbf{A}} \Rightarrow \varrho_1 = \varrho_2.$$

In a more general scheme one may perform several measurements and use all of their measurement outcome statistics. Let $\{\mathbf{A}, \mathbf{B}, \dots\}$ be a finite collection of POVMs. We denote by $\mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)$ the linear span of the set

$$\{\mathbf{A}_1, \dots, \mathbf{A}_n\} \cup \{\mathbf{B}_1, \dots, \mathbf{B}_m\} \cup \dots,$$

where n is the number of measurement outcomes of \mathbf{A} , m is the number of measurement outcomes of \mathbf{B} , and so forth. In particular,

$$\mathfrak{S}(\mathbf{A}) := \left\{ \sum_j c_j \mathbf{A}_j \mid c_j \in \mathbb{C} \right\}.$$

The set $\mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)$ is a linear subspace of $\mathcal{L}(\mathcal{H})$, the vector space of all linear operators on \mathcal{H} . Clearly, $\mathbb{1} \in \mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)$ and $\mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)^\dagger = \mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)$, where $\mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)^\dagger = \{A^\dagger : A \in \mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)\}$. A linear subspace of $\mathcal{L}(\mathcal{H})$ with these two properties is called an *operator system*.

If the system is described by a state ϱ and the measurement statistics of the collection $\{\mathbf{A}, \mathbf{B}, \dots\}$ is given, then we can calculate the expectation $\text{tr}[\varrho X]$ of any operator $X \in \mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)$. Therefore, from the point of view of informational completeness, only the linear span $\mathfrak{S}(\mathbf{A}, \mathbf{B}, \dots)$ is relevant. Two collections $\{\mathbf{A}, \mathbf{B}, \dots\}$ and $\{\mathbf{A}', \mathbf{B}', \dots\}$ are considered equivalent from this perspective if their linear spans are the same. The following simple result shows that any operator system is generated by a single POVM. Therefore, we can restrict our investigation to the situation where only a single POVM is measured.

Proposition 1 (POVMs and operator systems). *Let $\mathfrak{S} \subseteq \mathcal{L}(\mathcal{H})$ be an operator system (i.e. \mathfrak{S} is a linear subspace such that $\mathbb{1} \in \mathfrak{S}$ and $\mathfrak{S}^\dagger = \mathfrak{S}$). There exists a POVM \mathbf{A} such that $\mathfrak{S} = \mathfrak{S}(\mathbf{A})$ and \mathbf{A} has $\dim \mathfrak{S}$ outcomes. Any POVM \mathbf{B} satisfying $\mathfrak{S} = \mathfrak{S}(\mathbf{B})$ has at least $\dim \mathfrak{S}$ outcomes.*

Proof. Every operator $X \in \mathfrak{S}$ can be written as a linear combination of two selfadjoint operators,

$$X = \frac{1}{2}(X + X^\dagger) + i\frac{1}{2i}(X - X^\dagger),$$

which belong to \mathfrak{S} . If $X \in \mathfrak{S}$ is selfadjoint, then X can be written as a linear combination of two positive operators,

$$X = \frac{1}{2}(\|X\| \mathbb{1} + X) - \frac{1}{2}(\|X\| \mathbb{1} - X),$$

which again belong to \mathfrak{S} . Therefore, we can choose a (not necessarily orthonormal) basis for \mathfrak{S} consisting of positive operators and the identity $\mathbb{1}$; let $\{E_1, \dots, E_m, \mathbb{1}\}$ be such. For each $j = 1, \dots, m$, we set $\mathbf{A}_j = \frac{1}{m\|E_j\|} E_j$. The element \mathbf{A}_{m+1} is defined as $\mathbb{1} - \sum_{j=1}^m \mathbf{A}_j$. Then the collection $\mathbf{A}_1, \dots, \mathbf{A}_{m+1}$ forms a POVM and $\mathfrak{S} = \mathfrak{S}(\mathbf{A})$.

For any POVM \mathbf{B} , the generated operator system $\mathfrak{S}(\mathbf{B})$ cannot have more linearly independent elements than \mathbf{B} has outcomes. This implies the last claim. \square

We denote by \mathfrak{S}^\perp the orthogonal complement of an operator system \mathfrak{S} with respect to the Hilbert-Schmidt inner product. In particular, for a POVM $\mathbf{A} = \{\mathbf{A}_j\}_{j=1}^n$ we have

$$\mathfrak{S}(\mathbf{A})^\perp = \{B \in \mathcal{L}(\mathcal{H}) : \text{tr}[B^\dagger \mathbf{A}_j] = 0 \ \forall j = 1, \dots, n\}.$$

We notice that $\text{tr}[M] = 0$ for any $M \in \mathfrak{S}^\perp$ (since $\mathbb{1} \in \mathfrak{S}$) and $(\mathfrak{S}^\perp)^\dagger = \mathfrak{S}^\perp$ (since $\mathfrak{S}^\dagger = \mathfrak{S}$). The latter property implies that \mathfrak{S}^\dagger is spanned by its selfadjoint elements. Since $\mathfrak{S} \oplus \mathfrak{S}^\perp = \mathcal{L}(\mathcal{H})$, we have

$$\dim \mathfrak{S} + \dim \mathfrak{S}^\perp = d^2. \quad (1)$$

The complement $\mathfrak{S}(\mathbf{A})^\perp$ of $\mathfrak{S}(\mathbf{A})$ is related to the informational completeness of \mathbf{A} in the following simple way.

Proposition 2 (Operator systems and informational completeness). *Let $\mathfrak{S} \subseteq \mathcal{L}(\mathcal{H})$ be an operator system and let $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ be a set of states. Then a POVM \mathbf{A} satisfying $\mathfrak{S}(\mathbf{A}) = \mathfrak{S}$ is informationally complete w.r.t. \mathcal{P} iff*

$$(\mathcal{P} - \mathcal{P}) \cap \mathfrak{S}^\perp = \{0\}.$$

In other words, for any pair of states $\varrho_1, \varrho_2 \in \mathcal{S}(\mathcal{H})$ the following are equivalent:

- (i) $\varrho_1^\mathbf{A} = \varrho_2^\mathbf{A}$,
- (ii) $\varrho_1 - \varrho_2 \in \mathfrak{S}(\mathbf{A})^\perp$.

Proof. For all states ϱ_1, ϱ_2 , we have:

$$\begin{aligned} \varrho_1^\mathbf{A} = \varrho_2^\mathbf{A} &\iff \text{tr}[\varrho_1 \mathbf{A}] = \text{tr}[\varrho_2 \mathbf{A}] \quad \forall \mathbf{A} \in \mathfrak{S} \\ &\iff \text{tr}[(\varrho_1 - \varrho_2) \mathbf{A}] = 0 \quad \forall \mathbf{A} \in \mathfrak{S} \\ &\iff \varrho_1 - \varrho_2 \in \mathfrak{S}(\mathbf{A})^\perp. \end{aligned}$$

\square

For illustrative purposes, let us use the above framework to confirm the well known fact that there exists an informationally complete POVM with d^2 outcomes and that d^2 is the minimal number for which informational completeness can be achieved for all of $\mathcal{S}(\mathbb{C}^d)$ [11]. The set of all operators $\mathcal{L}(\mathbb{C}^d)$ is an operator system with dimension d^2 . By Proposition 1 there exists a POVM \mathbf{A} with d^2 outcomes and satisfying $\mathfrak{S}(\mathbf{A}) = \mathcal{L}(\mathbb{C}^d)$. Since $\mathfrak{S}(\mathbf{A})^\perp = \{0\}$, Proposition 2 implies that \mathbf{A} is informationally complete w.r.t. all states. On the other hand, if \mathbf{B} is a POVM with less than d^2 elements, then $\mathfrak{S}(\mathbf{B})^\perp$ contains a nonzero selfadjoint operator X (since $\mathfrak{S}(\mathbf{B})^\perp$ is generated by its selfadjoint part). The states $\frac{1}{d}\mathbb{1}$ and $\frac{1}{d}(\mathbb{1} + \|X\|^{-1}X)$ are different but they can not be distinguished by \mathbf{B} .

In some tomography schemes one may only infer from the expectation values of measurements rather than from the entire measurement outcome statistics. Such expectation values are characterized by a collection of selfadjoint operators. Let $\{S_1, S_2, \dots, S_n\}$ be a finite collection of selfadjoint operators. Again, if we know the expectations $\text{tr}[\varrho S_1], \dots, \text{tr}[\varrho S_n]$, we can calculate $\text{tr}[\varrho X]$ for any operator X belonging to the linear span of S_1, \dots, S_n .

Definition 2 (\mathcal{P} -informational completeness for selfadjoint operators). Let $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ be a subset of states. A collection $\{S_1, S_2, \dots, S_n\}$ of selfadjoint operators is called *informationally complete w.r.t. \mathcal{P}* iff

$$\text{tr}[\varrho_1 S_j] = \text{tr}[\varrho_2 S_j] \quad \forall j \Rightarrow \varrho_1 = \varrho_2$$

for all $\varrho_1, \varrho_2 \in \mathcal{P}$.

We are interested in questions of the type: what is the minimal number of

- 1) POVM outcomes
- 2) selfadjoint operators

needed to have an informationally complete measurement scheme w.r.t. to a given premise $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$?

Let us notice that the identity operator $\mathbb{1}$ does not give any useful information since $\text{tr}[\varrho \mathbb{1}] = 1$ for all states ϱ . Since an operator system $\mathfrak{S}(\mathbf{A})$ related to any POVM \mathbf{A} contains the identity operator $\mathbb{1}$, the minimal number of POVM outcomes is higher than the minimal number of selfadjoint operators. The above two questions 1) – 2) are thus related in the following way.

Proposition 3. *Let $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$. The following are equivalent.*

- (i) *There exists a POVM \mathbf{A} with n outcomes and \mathbf{A} is informationally complete w.r.t. \mathcal{P} .*
- (ii) *There exists a set $\{S_1, S_2, \dots, S_{n-1}\}$ of $n-1$ selfadjoint operators and $\{S_1, S_2, \dots, S_{n-1}\}$ is informationally complete w.r.t. \mathcal{P} .*

Proof. (i) \Rightarrow (ii): Suppose \mathbf{A} has n outcomes and \mathbf{A} is informationally complete w.r.t. \mathcal{P} . Since $\mathbf{A}_n = \mathbb{1} - \sum_{j=1}^{n-1} \mathbf{A}_j$, we conclude that the set $\{\mathbf{A}_1, \dots, \mathbf{A}_{n-1}\}$ of selfadjoint operators is informationally complete w.r.t. \mathcal{P} .

(ii) \Rightarrow (i): Suppose $\{S_1, S_2, \dots, S_{n-1}\}$ is a set of selfadjoint operators which is informationally complete w.r.t. \mathcal{P} . For each nonzero S_j , we define

$$\mathbf{A}_j := \left(\frac{1}{2} \mathbb{1} + \frac{1}{2} \|S_j\|^{-1} S_j \right) / (n-1).$$

Then $0 \leq \mathbf{A}_j \leq \mathbb{1}/(n-1)$ and by setting $\mathbf{A}_n := \mathbb{1} - \sum_{j=1}^{n-1} \mathbf{A}_j$ we obtain a POVM which is informationally complete w.r.t. \mathcal{P} . \square

As a consequence of Proposition 3 we only need to calculate the minimal number of either selfadjoint operators or POVM elements.

Definition 3. (Minimal informationally complete measurements) Let $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ be any subset of states. We denote by $\mathbf{m}[\mathcal{P}]$

- the minimal number of selfadjoint operators which are informationally complete w.r.t. \mathcal{P} .

or, equivalently (by Prop. 3),

- the minimal number of positive operators determining a POVM which is informationally complete w.r.t. \mathcal{P} (i.e., the number of POVM elements subtracted by one).

The remaining part of this work is devoted to deriving upper and lower bounds on $\mathbf{m}[\mathcal{P}]$ either for various concrete subsets \mathcal{P} or depending on specific properties of \mathcal{P} – in particular its dimensionality and topology.

3. UPPER BOUNDS

3.1. States with bounded rank. As the first class of examples, we consider the subset $\mathcal{P}_r := \{\varrho \in \mathcal{S}(\mathcal{H}) : \text{rank}(\varrho) \leq r\}$ of all states whose rank is bounded by a number $r < d/2$ where $d = \dim \mathcal{H}$. How many measurement outcomes suffice in order to identify a state taken from \mathcal{P}_r ?

Proposition 4 (States with bounded rank). *If $1 \leq r < d/2$, then there exists a POVM \mathbf{A} which is informationally complete w.r.t. \mathcal{P}_r and has $4r(d-r)$ outcomes. Therefore,*

$$\mathfrak{m}[\mathcal{P}_r] \leq 4r(d-r) - 1.$$

Proof. We will construct a subspace \mathcal{B} of $d \times d$ matrices with the properties that

- (a) $\mathcal{B}^\dagger = \mathcal{B}$,
- (b) $\text{tr}[B] = 0$ for every $B \in \mathcal{B}$,
- (c) $\dim \mathcal{B} = (d-2r)^2$,
- (d) $\text{rank}(B) \geq 2r+1$ for every nonzero $B \in \mathcal{B}$.

From (a) – (b) it follows that $\mathfrak{S} := \mathcal{B}^\perp$ is an operator system. By Prop. 1 there exists a POVM \mathbf{A} such that $\mathfrak{S}(\mathbf{A}) = \mathfrak{S}$ and \mathbf{A} has $\dim \mathfrak{S}$ outcomes. From (c) we get that \mathbf{A} has $d^2 - (d-2r)^2 = 4r(d-r)$ outcomes. Finally, if $\varrho_1, \varrho_2 \in \mathcal{P}_r$, then $\text{rank}(\varrho_1 - \varrho_2) \leq 2r$. By Prop. 2 this implies that an observable \mathbf{A} is informationally complete w.r.t. \mathcal{P}_r if $\text{rank}(B) \geq 2r+1$ for every nonzero operator $B \in \mathfrak{S}(\mathbf{A})^\perp$. This is guaranteed by (d). Hence, constructing a subspace \mathcal{B} with the properties (a) – (d) will prove the proposition.

The main part of our construction follows [12]. The following fact will be needed. Let M be a totally nonsingular $m \times m$ -matrix with positive entries. (Recall that a matrix is called totally nonsingular if all of its minors are nonzero.) For instance, a Vandermonde matrix of the form

$$M = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{m-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_m & \alpha_m^2 & \cdots & \alpha_m^{m-1} \end{pmatrix}$$

with $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_m$ has strictly positive minors [13]. Since M is totally nonsingular, any linear combination of ℓ columns of M contains at most $\ell - 1$ zero elements: if some linear combination of ℓ columns of M would contain ℓ zero elements, then we could find a $\ell \times \ell$ submatrix of M with linearly dependent columns. This contradicts the requirement that all the minors of M are nonzero.

We will now define a set of $d \times d$ matrices that span the sought subspace \mathcal{B} . By the k th diagonal of a matrix $[M_{ij}]$ we mean the elements M_{ij} with $i - j = d - k$. (In other words, we label the diagonals from the lower left corner upwards. The main diagonal is then the d th diagonal.)

For each k satisfying $2r+1 \leq k \leq d-1$, we build up $k-2r$ matrices as follows. We choose $k-2r$ columns from a totally nonsingular $k \times k$ -matrix and we put them to the k th diagonal and 0's elsewhere, hence obtaining $k-2r$ linearly independent matrices. Any matrix P which is a linear combination of these $k-2r$ matrices has at least $2r+1$ nonzero elements on the k th diagonal. Since all the matrix elements not in the k th diagonal are 0, we see that the P has rank at least $2r+1$.

We also take all transposes of the previously constructed matrices to our spanning set of matrices. In addition we construct and add to the set $d - 2r$ diagonal matrices (with nonzero entries only on the $k=d$ th diagonal). To this end, we choose $d - 2r$ columns from a totally nonsingular $d \times d$ -matrix M . Let v_1, \dots, v_{d-2r} denote these column vectors. Next, we choose a real vector u which has no zero entries and which is orthogonal to every v_1, \dots, v_{d-2r} . A possible choice is, for instance, the last row from the inverse matrix of M . Then the new vectors $\tilde{v}_1, \dots, \tilde{v}_{d-2r}$ defined as entrywise products of v_1, \dots, v_{d-2r} with u each have the property that their components sum up to zero. Hence, from these vectors we can construct $d - 2r$ traceless diagonal matrices such that again any non-zero linear combination of them has rank at least $2r + 1$.

In total, we have built up $d - 2r + 2 \sum_{k=2r+1}^{d-1} (k - 2r) = (d - 2r)^2$ linearly independent matrices. The previously mentioned subspace \mathcal{B} is the linear span of these matrices. The properties (a)-(c) are immediate consequences of the construction. To verify (d), suppose that $B \in \mathcal{B}$. Let k_B be the largest k such that the k th diagonal of B contains nonzero elements. Then the k_B th diagonal contains actually $2r + 1$ nonzero elements. The $(2r + 1) \times (2r + 1)$ submatrix having those $2r + 1$ nonzero elements in its main diagonal is lower triangular, therefore has nonzero determinant. It follows that the rank of B is at least $2r + 1$. \square

Let us remark that a recently introduced method based on compressed sensing uses $O(dr \log(d)^2)$ outcomes to identify a rank- r state with high probability [6]. The number of outcomes given in Proposition 4 therefore beats the compressed sensing approach. The latter, however, might be advantageous regarding the simplicity of the classical post-processing, robustness and verifiability of the assumption.

3.2. Pure states. We will now have a closer look at the set \mathcal{P}_1 of pure states. For this case, Proposition 4 implies that $\mathfrak{m}[\mathcal{P}_1] \leq 4d - 5$. We will first provide some simple arguments showing that indeed $\mathfrak{m}[\mathcal{P}_1] = 4d - 5$ for $d = 2, 3$. In general, we will see later, based on topological reasoning, that the leading order $4d$ is the best possible.

We remark that Flammia et al. proved that a POVM which is informationally complete w.r.t. the set of pure states has at least $2d$ outcomes [5]. They also constructed a POVM with $2d$ elements capable of distinguishing almost all (but not all) pairs of pure states. A similar construction was given by Finkelstein [4].

Proposition 5. *For a POVM \mathbf{A} , the following conditions are equivalent:*

- (i) \mathbf{A} is not informationally complete w.r.t. the set of pure states.
- (ii) $\mathfrak{S}(\mathbf{A})^\perp$ contains a selfadjoint operator $T \neq 0$ with $\text{rank}(T) \leq 2$.
- (iii) $\mathfrak{S}(\mathbf{A})^\perp$ contains a selfadjoint operator T with $\text{rank}(T) = 2$.

Proof. (i) \Rightarrow (iii): Suppose \mathbf{A} is not informationally complete w.r.t. pure states. By Prop. 2 there exists two pure states $\varrho_1 \neq \varrho_2$ such that $\varrho_1 - \varrho_2 =: T \in \mathfrak{S}(\mathbf{A})^\perp$. The operator T is selfadjoint and $\text{rank}(T) = 2$.

(ii) \Leftrightarrow (iii) is seen by observing that there is no traceless selfadjoint operator with rank 1.

(iii) \Rightarrow (i) Suppose there is a selfadjoint operator $T \in \mathfrak{S}(\mathbf{A})^\perp$ with rank 2. Since $\text{tr}[T] = 0$ it has two nonzero eigenvalues $\pm\lambda$. The operator $T' := \frac{1}{\lambda}T$ then has spectral decomposition $T' = P_1 - P_2$, where P_1 and P_2 are one-dimensional projections, hence pure states. By Prop. 2 the POVM \mathbf{A} cannot distinguish P_1 and P_2 . \square

From Proposition 5 we conclude the following simple characterization.

Corollary 1 (Pure state informationally complete measurements). *A POVM \mathbf{A} is informationally complete w.r.t. the set of pure states if and only if every nonzero selfadjoint operator $T \in \mathfrak{S}(\mathbf{A})^\perp$ has $\text{rank}(T) \geq 3$.*

Remark 1 (The qubit case). It is an immediate consequence of Corollary 1 that in the qubit case (i.e. $\dim \mathcal{H} = 2$) a POVM \mathbf{A} is informationally complete w.r.t. pure states if and only if $\mathfrak{S}(\mathbf{A})^\perp = \{0\}$. Therefore, in the qubit case informational completeness for pure states implies informational completeness for all states. One can also easily see this by a direct inspection of the Bloch sphere.

Note that if $\mathfrak{S}(\mathbf{A})^\perp$ contains an operator M with $\text{rank}(M) = 1$, then $M + M^\dagger$ is selfadjoint and $\text{rank}(M + M^\dagger) \leq 2$. Thus, \mathbf{A} is not informationally complete w.r.t. pure states if $\mathfrak{S}(\mathbf{A})^\perp$ contains a rank-1 operator. On the other hand, if $\mathfrak{S}(\mathbf{A})^\perp$ contains an operator M with $\text{rank}(M) = 2$, this does *not* imply that there exists a selfadjoint operator T with $\text{rank}(T) = 2$. For instance, the subspace

$$\left\{ \begin{pmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha \\ \beta & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \end{pmatrix} : \alpha, \beta \in \mathbb{C} \right\}$$

contains rank-2 matrices but no rank-2 selfadjoint matrix.

As an application of Corollary 1, we can easily characterize all POVMs in dimension 3 which are informationally complete w.r.t. the set of pure states.

Proposition 6 (Pure state informational completeness in dimension 3). *A POVM \mathbf{A} is informationally complete w.r.t. the set of all pure states in $\mathcal{S}(\mathbb{C}^3)$ if and only if it falls into one of the two classes:*

- $\mathfrak{S}(\mathbf{A})^\perp = \{0\}$ (i.e. \mathbf{A} is informationally complete w.r.t. all states).
- $\mathfrak{S}(\mathbf{A})^\perp = \{cT : c \in \mathbb{C}\}$ for some invertible selfadjoint operator T with $\text{tr}[T] = 0$.

In particular, a minimal POVM which is informationally complete w.r.t. pure states has 8 outcomes.

Proof. We first prove the following: if $\dim \mathfrak{S}(\mathbf{A}) \leq 7$, then \mathbf{A} cannot distinguish all pairs of distinct pure states. We need to show that there exists a singular selfadjoint operator $T \in \mathfrak{S}(\mathbf{A})^\perp$. The claim then follows from Cor. 1. From $\dim \mathfrak{S}(\mathbf{A}) \leq 7$ follows that $\dim \mathfrak{S}(\mathbf{A})^\perp \geq 2$. There thus exist two linearly independent selfadjoint operators $X, Y \in \mathfrak{S}(\mathbf{A})^\perp$. If either X or Y is singular, we are done. So let us assume that $\det(X) > 0$ and $\det(Y) < 0$ (otherwise we redefine $X \rightarrow -X$ or $Y \rightarrow -Y$). By the intermediate value theorem the function $t \in \mathbb{R} \mapsto \det(tX + (1-t)Y)$ has to be zero for some $t = t_0 \in (0, 1)$. The corresponding operator $t_0X + (1-t_0)Y$ is thus a singular element of $\mathfrak{S}(\mathbf{A})^\perp$.

We conclude that if \mathbf{A} is informationally complete for the set of all pure states, then either $\dim \mathfrak{S}(\mathbf{A}) = 8$ or $\dim \mathfrak{S}(\mathbf{A}) = 9$. In the latter case \mathbf{A} is informationally complete for all states. If $\dim \mathfrak{S}(\mathbf{A}) = 8$, then $\dim \mathfrak{S}(\mathbf{A})^\perp = 1$ and $\mathfrak{S}(\mathbf{A})^\perp$ is thus generated by a single selfadjoint operator T . From Cor. 1 follows that \mathbf{A} is informationally complete w.r.t. all pure states exactly when T is invertible. \square

Proposition 6 shows how to construct POVMs that are informationally complete w.r.t. pure states for $d = 3$: fix any invertible selfadjoint operator T with $\text{tr}[T] = 0$. Then the complement space is an operator system from which the sought POVM can be obtained as a spanning set (see the proof of Proposition 1).

Remark 2 (Necessity of prior information). The characterization given in Proposition 6 helps to illustrate one possible drawback of POVMs which are merely informationally complete w.r.t. to a subset of states: even if a POVM A can distinguish any pair of distinct pure states, it may not be capable to distinguish pure states from mixed states. In other words, we may not be able to verify the premise from the measurement outcome statistics.

To give an example, let A be a POVM in \mathbb{C}^3 with $\mathfrak{S}(A)^\perp = \{cT : c \in \mathbb{C}\}$ for some invertible selfadjoint operator T with $\text{tr}[T] = 0$. Suppose T has a single positive eigenvalue (if not, take $-T$). Let $T = \lambda_1 P_1 - (\lambda_2 P_2 + \lambda_3 P_3)$, $\lambda_j > 0$, be the spectral decomposition of T into one-dimensional projections P_j . The POVM A cannot distinguish the pure state P_1 from the mixed state $\varrho = \lambda_2/\lambda_1 P_2 + \lambda_3/\lambda_1 P_3$. We conclude that by measuring A one cannot identify the pure state P_1 from the measurement outcome statistics *without* making use of the prior knowledge that the state is pure.

We already noted that Proposition 4 leads to the general upper bound $\mathfrak{m}[\mathcal{P}_1] \leq 4d - 5$ albeit without yielding an explicit set of operators. Now, we give another proof of this result via a simple direct construction. The construction is inspired by a topological embedding given by James in [14]. Consider two types of matrices X_α and Y_β which we label by integers $\alpha = 1, \dots, 2d - 2$ and $\beta = 1, \dots, 2d - 3$ respectively. The X_α 's are taken to be such that $(X_\alpha)_{kl} = \delta_{k+l, \alpha+1}$, i.e., there are 1's along the α 'th anti-diagonal and zeroes elsewhere. The Y_β 's are similarly defined with nonzero entries solely along the anti-diagonals, in this case $(Y_\beta)_{kl} = 0$ unless $k + l = \beta + 2$. The entries are chosen such that the matrices are anti-symmetric with entries i below the diagonal. The constructed matrices thus have the following structure:

$$X \sim \begin{pmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{pmatrix}, \quad \text{and} \quad Y \sim \begin{pmatrix} & & & -i \\ & & -i & \\ & i & & \\ i & & & \end{pmatrix}. \quad (2)$$

Proposition 7 (Pure state informational completeness—explicit construction). *The set $\{S_j\} := \{X_\alpha, Y_\beta\}$ consisting of $4d - 5$ selfadjoint operators is informationally complete w.r.t. the set of all pure states on \mathbb{C}^d .*

Proof. Following [14] we use the following type of auxiliary matrices. Consider a set of upper-triangular matrices $\{C_\gamma \in \mathbb{C}^{d \times d}\}_{\gamma=2, \dots, 2d}$ which are such that $(C_\gamma)_{kl} = 0$ if $k + l > \gamma$ and $(C_\gamma)_{kl} \neq 0$ if $k + l = \gamma$. We claim that for all vectors $x, y \in \mathbb{C}^d$ the following holds:

$$\begin{aligned} &\text{if } \langle x | C_\gamma | x \rangle = \langle y | C_\gamma | y \rangle \text{ for all } \gamma, \\ &\text{then } y = e^{i\varphi} x \text{ for some } \varphi \in \mathbb{R}. \end{aligned} \quad (3)$$

To prove this claim, assume that the n 'th component of x is the first with a non-zero entry $x_n \neq 0$. Then $\langle x | C_{2n} | x \rangle = (C_{2n})_{n,n} |x_n|^2$ which by the hypothesis in Eq.(3) implies that there is a $\varphi \in \mathbb{R}$ so that $y_l = e^{i\varphi} x_l$ holds for all $l \leq n$ (since $x_l = y_l = 0$ for $l < n$ and

$|x_n| = |y_n|$). By induction we can now prove that the same proportionality has to hold for all other components. So assume for some $m \geq n$ that $y_l = e^{i\varphi} x_l$ holds for all $l \leq m$. Then for $\gamma = m + n + 1$

$$\begin{aligned} \langle x | C_\gamma | x \rangle &= \sum_{i+j \leq \gamma} (C_\gamma)_{ij} \bar{x}_i x_j \\ &= (C_\gamma)_{n,m+1} e^{i\varphi} \bar{y}_n x_{m+1} + \sum_{i+j \leq m+n} (C_\gamma)_{ij} \bar{y}_i y_j, \end{aligned} \quad (4)$$

where we replaced $x_l \rightarrow e^{-i\varphi} y_l$ for all $l \leq m$ and exploited that C_γ is upper triangular and that $x_l = 0$ for all $l < n$. Together with the hypothesis in Eq.(3) and the assumption that $(C_\gamma)_{n,m+1} \neq 0$ this implies indeed that $y_{m+1} = e^{i\varphi} x_{m+1}$.

We now exploit Eq.(3) for specific matrices which we construct as $C_{2d} = \mathbb{1}$ and for $\gamma = 2, \dots, 2d - 1$ as

$$(C_\gamma)_{kl} = \begin{cases} \delta_{k+l,\gamma}, & k < l, \\ 1/2, & k = l = \gamma/2 \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Note that $C_\gamma = (X_{\gamma-1} + iY_{\gamma-2})/2$ for $\gamma = 3, \dots, 2d - 1$ and $C_2 = X_1/2$. Hence, if $\langle x | S_j | x \rangle = \langle y | S_j | y \rangle$ for every j , then $\langle x | C_\gamma | x \rangle = \langle y | C_\gamma | y \rangle$ for every $\gamma = 2, \dots, 2d - 1$. The remaining condition $\langle x | C_{2d} | x \rangle = \langle y | C_{2d} | y \rangle$ holds due to $\|x\| = \|y\|$. Therefore it follows from Eq.(3) that the set $\{S_j\}$ is informationally complete w.r.t. the set of all pure states. \square

It is possible to slightly improve the upper bound $\mathbf{m}[\mathcal{P}_1] \leq 4d - 5$ at the cost of having to deal with a more complicated set of operators introduced in the work of Milgram [15]:

Proposition 8 (Pure state informational completeness – improved bound). *Let α denote the number of 1's in the binary expansion of $d - 1$. There exists a collection of m selfadjoint operators which is informationally complete w.r.t. pure states in \mathbb{C}^d , if*

$$m = \begin{cases} 4d - 4 - \alpha & \text{for odd } d, \\ 4d - 5 - \alpha & \text{for even } d \geq 4. \end{cases} \quad (6)$$

The improvement of this result to our earlier bound $\mathbf{m}[\mathcal{P}] \leq 4d - 5$ is thus α (for even $d \geq 4$) or $\alpha - 1$ (for odd d). Notice that α satisfies $1 \leq \alpha \leq \log_2(d)$. We will see later that this improvement is, in fact, nearly optimal.

Proof. The construction of the operators is based on the work of Milgram [15]. We will only argue why this corresponds to a proper measurement scheme rather than reproducing the rather cumbersome construction.

For m as in the proposition, Milgram constructed a set of m bilinear maps, i.e., matrices $A_j \in \mathbb{C}^{d \times d}$, $j = 1, \dots, m$ which have the following properties:

(i) *Vanishing real inner product* in the sense that for all $x \in \mathbb{C}^d$ we have $\langle x, A_j x \rangle_{\mathbb{R}} = 0$ for the real inner product $\langle x, y \rangle_{\mathbb{R}} = \Re \langle x, y \rangle$. That is, each A_j has to be skew-symmetric w.r.t. the real inner product and thus anti-selfadjoint w.r.t. to the standard complex inner product. In order to see the latter note that $\langle x, A_j y \rangle_{\mathbb{R}} + \langle y, A_j x \rangle_{\mathbb{R}} = 0 \ \forall x, y \in \mathbb{C}^d$ can be written as

$$\Re \langle \psi | \begin{pmatrix} 0 & A_j \\ A_j & 0 \end{pmatrix} | \psi \rangle = 0, \quad \forall \psi \in \mathbb{C}^{2d}. \quad (7)$$

Hence, the selfadjoint part of the anti-diagonal block matrix has to vanish which indeed means that $A_j^\dagger = -A_j$.

(ii) *Completeness.* We can define matrices $T_j = iA_j$, which according to (i) are selfadjoint, and as noted by Mukherjee [16] such that the map $f : \mathbb{C}^d \rightarrow \mathbb{R}^m$ defined via $f(x)_j = \langle x, T_j x \rangle$ has the property that $f(x) = f(y)$ implies that x is proportional to y .

The set of m selfadjoint matrices T_j therefore leads to a measurement scheme which is informationally complete w.r.t. the set of pure states. \square

3.3. Generic bounds depending on fractal dimension. In this part we will discuss general upper bounds on $\mathfrak{m}[\mathcal{P}]$ which depend only on the dimensionality of \mathcal{P} . To this end, it is useful to regard any measurement scheme as a linear map between real Euclidean vector spaces. More precisely, we identify the set of selfadjoint operators on \mathbb{C}^d with \mathbb{R}^{d^2} and the set of m -tuples of selfadjoint operators in $\mathcal{L}(\mathbb{C}^d)$ as linear maps from \mathbb{R}^{d^2} into \mathbb{R}^m .

Let us assume that \mathcal{P} is a *closed* subset of $\mathcal{S}(\mathcal{H})$. Then \mathcal{P} can be identified with a compact subset of \mathbb{R}^{d^2} . The *Minkowski dimension* (also called *box dimension*) $D(\mathcal{P})$ of \mathcal{P} is obtained by considering the minimal number $N_\epsilon(\mathcal{P})$ of ϵ -balls needed to cover \mathcal{P} and taking the limit

$$D(\mathcal{P}) := \limsup_{\epsilon \rightarrow 0} \frac{\log(N_\epsilon(\mathcal{P}))}{\log(1/\epsilon)}.$$

For instance, if \mathcal{P} is a smooth manifold of real dimension $d(\mathcal{P})$, then $D(\mathcal{P}) = d(\mathcal{P})$.

By Mané's theorem [17, 18] for a compact set \mathcal{P} almost any (in the Lebesgue measure sense) linear map Λ from \mathbb{R}^{d^2} into \mathbb{R}^m is injective on \mathcal{P} if $m > 2D(\mathcal{P})$. The injectivity of Λ is clearly equivalent to the property that the corresponding set of m selfadjoint operators is informational complete w.r.t. \mathcal{P} . Hence, we conclude the following result.

Proposition 9 (Informational completeness of generic measurements). *Let $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ be a closed subset of the set of density matrices with Minkowski dimension $D(\mathcal{P})$. Then almost any collection of $m > 2D(\mathcal{P})$ selfadjoint operators is informationally complete w.r.t. \mathcal{P} . In particular, $\mathfrak{m}[\mathcal{P}] \leq 2D(\mathcal{P}) + 1$.*

Remark 3. In principle, this bound can be refined to $m > \delta_{\mathcal{P}-\mathcal{P}}$, where $\delta_{\mathcal{P}-\mathcal{P}}$ is the *Hausdorff dimension* of the set $\mathcal{P} - \mathcal{P} := \{\varrho_1 - \varrho_2 | \varrho_i \in \mathcal{P}\}$ [19]. This bound is generally better since $\delta_{\mathcal{P}-\mathcal{P}} \leq D(\mathcal{P} - \mathcal{P}) \leq 2D(\mathcal{P})$, but it may be more difficult to handle. We also mention that for the inverse mappings Hölder continuity can be proven and the respective constants can be bounded [20, 18].

Let us apply Proposition 9 to the case of pure states. This is a smooth manifold of real dimension $2d - 2$, hence $D(\mathcal{P}_1) = 2d - 2$. Thus, almost any collection of $4d - 3$ selfadjoint operators is informationally complete w.r.t. pure states.

A related example for which it might be difficult to obtain good bounds by other means is the set of depolarized pure states. Let $\sigma \in \mathcal{S}(\mathbb{C}^d)$ be any mixed state and define $\mathcal{P}_\sigma := \{\varrho \in \mathcal{S}(\mathbb{C}^d) | \varrho \in \lambda\sigma + (1 - \lambda)\mathcal{P}_1, \lambda \in [0, 1]\}$. Then $D(\mathcal{P}_\sigma) = D(\mathcal{P}_1) + 1$ so that almost any set of $4d - 1$ selfadjoint operators is informationally complete w.r.t. \mathcal{P}_σ .

4. LOWER BOUNDS

The main idea which in this section provides lower bounds on \mathfrak{m} is, that an informationally complete measurement preserves the topology of \mathcal{P} when we regard the measurement

scheme as a mapping from \mathcal{P} into the real vector space \mathbb{R}^m corresponding to measurement outcomes or probabilities. More precisely, on its image in \mathbb{R}^m this map has to be a homeomorphism or, with some additional assumptions, a diffeomorphism. For various manifolds \mathcal{P} the existence of such maps, i.e., the possibility of a (differential) topological embedding into \mathbb{R}^m is well studied and allows us to translate non-embedding results into lower bounds on m . Fortunately, in some cases of interest these bounds are very close to or even match the upper bounds obtained in the previous section.

4.1. Measurements as homeomorphism. As before we will identify the set of selfadjoint operators on \mathbb{C}^d with \mathbb{R}^{d^2} . In this way, we can regard the set $\mathcal{S}(\mathcal{H})$ of all states as well as any of its closed subsets $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ as a compact subset of \mathbb{R}^{d^2} from which it inherits a natural topology.

Let \mathbf{A} be a POVM with $m + 1$ outcomes. This induces a map

$$h_{\mathbf{A}} : \mathcal{S}(\mathcal{H}) \rightarrow \mathbb{R}^m, \quad h_{\mathbf{A}}(\varrho) := (\text{tr}[\varrho \mathbf{A}_1], \dots, \text{tr}[\varrho \mathbf{A}_m]), \quad (8)$$

which is injective as a map from \mathcal{P} iff \mathbf{A} is informationally complete w.r.t. \mathcal{P} . With a slight abuse of notation we write $h_{\mathbf{A}}$ for the map from \mathcal{P} as well as for the extended map from \mathbb{R}^{d^2} .

Proposition 10 (Informational completeness and topological embeddings). *Let $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ be a closed subset. A POVM \mathbf{A} is informationally complete w.r.t. \mathcal{P} iff the map $h_{\mathbf{A}}$ is a topological embedding of \mathcal{P} into \mathbb{R}^m .*

Proof. By definition a topological embedding is an injective continuous map which has a continuous inverse on its image, i.e., a homeomorphism onto its image. So injectivity of $h_{\mathbf{A}}$ and thus informational completeness of \mathbf{A} is implied by $h_{\mathbf{A}}$ being a topological embedding. For the converse note that $h_{\mathbf{A}}$ is linear on \mathbb{R}^{d^2} and thus continuous. Moreover, by assumption $h_{\mathbf{A}} : \mathcal{P} \rightarrow h_{\mathbf{A}}(\mathcal{P})$ is a continuous bijection and as such has a continuous inverse since \mathcal{P} is compact. \square

The usefulness of Proposition 10 is that it gives a method to derive a lower bound for $m[\mathcal{P}]$. If we know that \mathcal{P} does not admit a topological embedding into \mathbb{R}^m , then $m[\mathcal{P}] > m$.

Remark 4 (Pure qubit states revisited). As a simple demonstration of Proposition 10, let us consider the subset of all pure qubit states. This subset is homeomorphic to the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. Hence, any POVM with 3 outcomes defines a continuous map from \mathbb{S}^2 into \mathbb{R}^2 . A measurement separating all pairs of distinct pure states has to map different initial states (=points on the sphere) to different points in \mathbb{R}^2 . A discontinuous mapping (option (b) in Fig.1) cannot arise from a measurement. Every continuous map into the plane, however, necessarily identifies different points of the sphere (option (a) in Fig.1). In fact, by the *Borsuk-Ulam theorem* every continuous map from \mathbb{S}^{n-1} into \mathbb{R}^n maps some pair of antipodal points to the same point. Consequently, we recover the fact that any pure state informationally complete POVM on a qubit must have more than 3 outcomes.

Remark 5 (Measurements on several copies). Although we do not discuss this scenario in other parts of this work, let us mention that Proposition 10 holds as well if we allow global measurements on multiple copies, i.e., if we replace ϱ by $\varrho^{\otimes n}$ for some $n \in \mathbb{N}$. The map $h_{\mathbf{A}}$ then becomes non-linear but it is evidently still continuous, which is all we need. In this way, also the aforementioned qubit results generalize to the multiple copy case.

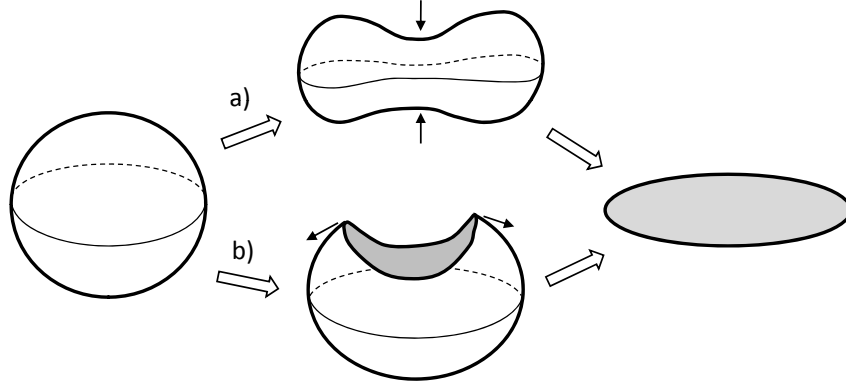


FIGURE 1. A mapping from the sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ to the plane \mathbb{R}^2 is either (a) not injective, or (b) not continuous. When regarding the sphere as the manifold of pure qubit states and the mapping as a measurement with 3 outcomes (of which only two can have independent probabilities), this simple picture implies that a pure state informationally complete measurement requires 4 outcomes.

By a *surface* we mean a 2-dimensional topological manifold, i.e., a topological space with the property that every point has a neighborhood which is homeomorphic to an open subset of \mathbb{R}^2 . A surface is orientable or non-orientable. (We recall that a possible characterization of this distinction is that a surface is non-orientable if and only if it contains a homeomorphic image of the Möbius strip.)

Corollary 2 (2-manifolds). *If $\mathcal{P} \subset \mathcal{S}(\mathcal{H})$ is a closed surface without boundary, then $\mathfrak{m}[\mathcal{P}] \geq 3$. Moreover, if \mathcal{P} is non-orientable in addition, then $\mathfrak{m}[\mathcal{P}] \geq 4$.*

Proof. This follows from Prop.10 by noting that (i) a compact surface without boundary cannot be embedded into \mathbb{R}^2 (where any compact set does have a boundary) and (ii) by the classification of surfaces, non-orientability does not permit an embedding into \mathbb{R}^3 . \square

We note that this type of purely topological reasoning cannot give better lower bounds since by Whitney's embedding theorem every surface can be embedded in \mathbb{R}^4 [21].

As an application of Corollary 2, let us consider the set of pure states in \mathbb{C}^3 with real amplitudes. That is, we fix an orthonormal basis $\{\varphi_j\}_{j=1}^3$ and the pure states under investigation correspond to the vectors $\psi = \sum_{j=1}^3 r_j \varphi_j$, $r_j \in \mathbb{R}$.

Corollary 3 (Pure states with real amplitudes in \mathbb{C}^3). *Let $\mathcal{P} \subset \mathcal{S}(\mathbb{C}^3)$ be the set of pure states with real amplitudes w.r.t. a fixed basis in \mathbb{C}^3 . Then $\mathfrak{m}[\mathcal{P}] = 4$.*

Proof. Due to normalization every state in \mathcal{P} can be represented by a unit vector $x \in \mathbb{S}^2$. Since x and $-x$, however, represent the same state, we have to identify antipodes so that \mathcal{P} is homeomorphic with \mathbb{RP}^2 , the *real projective plane*. Since \mathbb{RP}^2 is non-orientable, it follows from Cor. 2 that $\mathfrak{m}[\mathbb{RP}^2] \geq 4$. In order to see that in fact $\mathfrak{m}[\mathbb{RP}^2] = 4$ consider the map $x \mapsto (x_1 x_2, x_2 x_3, x_3 x_1, x_1^2 - x_2^2)$. This is a topological embedding of \mathbb{RP}^2 in \mathbb{R}^4 which can be realized by a measurement scheme: the first three components can be obtained by σ_x -type measurements and the fourth via a σ_z -type measurement. \square

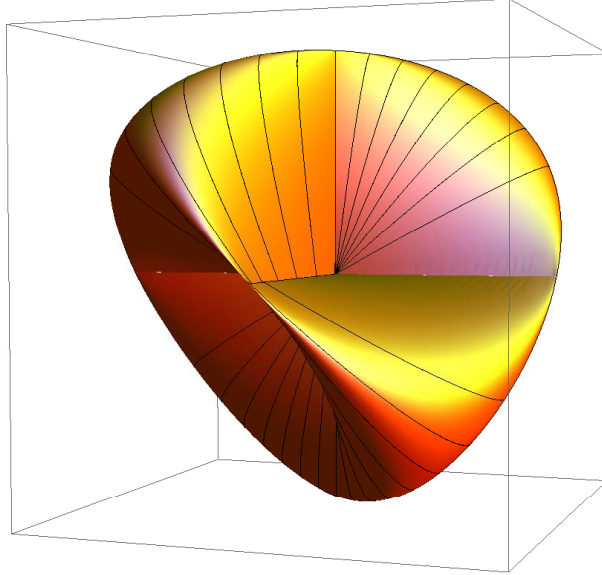


FIGURE 2. The *Roman surface* is a self-intersecting surface in \mathbb{R}^3 obtained by a mapping from the real projective plane \mathbb{RP}^2 . In our context it arises when the manifold of three-dimensional real, pure states is mapped onto the expectation values of three σ_x -type observables. Since the manifold is not orientable, every three-outcome measurement has to be non-injective—here reflected by the self-intersections of the surface.

Remark 6 (Roman surface). The three σ_x -type measurements in the proof of Corollary 3 give rise to measurement results which form the *Roman surface* displayed in Fig.2. The failure of informational completeness (due to disregarding the necessary fourth measurement) is reflected by self-intersections of the surface.

4.2. Measurements as diffeomorphisms. Manifolds of interest in quantum tomography often arise from unitary orbits, so that they have a differentiable structure – they are *smooth manifolds*. In such a case we may resort to differential topology which imposes more restrictive conditions on the existence of *smooth embeddings*. Before we apply these to the concrete cases of pure states and states with general rank constraints we provide some general background.

Suppose that \mathcal{P} is a smooth manifold. A smooth embedding f of \mathcal{P} into \mathbb{R}^m is a smooth map which is a topological embedding (i.e. homeomorphism onto its image) and has the property that the derivative of f is everywhere injective. Although for most smooth manifolds the minimal embedding dimension m is not known exactly, quite narrow intervals have been determined for many cases of interest (see [22, 23] for an overview). One tool to derive lower bounds on the minimal embedding dimension m is Chern’s result [24] that a smooth embedding into \mathbb{R}^m requires that the dual Stiefel-Whitney classes vanish $\bar{W}(\mathcal{P})_i = 0$ for all $i \geq m - D(\mathcal{P})$. Other bounds can be obtained from an index theorem due to Atiyah and Hirzebruch [25] and similar ideas [26, 27]. On the positive side a general upper bound is due to Whitney [21] who showed that a smooth embedding $\mathcal{P} \rightarrow \mathbb{R}^m$ always exists if $m \geq 2D(\mathcal{P})$ (actually $m \geq 2D(\mathcal{P}) - 1$ unless m is a power of 2). Whitney’s bound is known to be optimal,

i.e., in the worst case the dimension of the Euclidean space has to be twice the dimension of the manifold.

Again, for a POVM \mathbf{A} with $m + 1$ outcomes, we denote by $h_{\mathbf{A}} : \mathcal{S}(\mathcal{H}) \rightarrow \mathbb{R}^m$ the induced mapping defined in Eq.(8). In order to apply the known results on the lower bounds on dimensions for smooth embeddings, we have to show that for a \mathcal{P} -informationally complete POVM \mathbf{A} , the related map $h_{\mathbf{A}} : \mathcal{P} \rightarrow h_{\mathbf{A}}(\mathcal{P})$ is a diffeomorphism. In Subsection 4.1 we saw already that it is a homeomorphism.

Throughout we will suppose that \mathcal{P} is a *compact embedded submanifold* of \mathbb{R}^{d^2} where we identify the latter with the space of selfadjoint matrices in $\mathbb{C}^{d \times d}$. With a slight abuse of notation we will write \mathcal{P} for both, the manifold and its inclusion in \mathbb{R}^{d^2} . We denote by $T_p(\mathcal{P})$ the tangent space of \mathcal{P} at $p \in \mathcal{P}$ and by $h_* : T_p(\mathcal{P}) \rightarrow T_{h_{\mathbf{A}}(p)}(h(\mathcal{M}))$ the derivative, which is a linear map between the tangent spaces (sometimes call pushforward). The cone defined by

$$\Delta(\mathcal{P}) := \{X \in \mathbb{R}^{d^2} | X = \lambda(M_1 - M_2) \text{ for some } M_i \in \mathcal{P}, \lambda > 0\}.$$

will play an important role in the following.

Proposition 11 (Smooth embeddings). *Let \mathcal{P} be a compact embedded submanifold of \mathbb{R}^{d^2} , where the latter is identified with the space of selfadjoint matrices in $\mathbb{C}^{d \times d}$. Suppose that $h_{\mathbf{A}} : \mathcal{P} \rightarrow \mathbb{R}^m$ is a mapping of the form in Eq.(8). Then $h_{\mathbf{A}}(\mathcal{P})$ is a smooth embedding of \mathcal{P} in \mathbb{R}^m if $h_{\mathbf{A}}$ is injective on \mathcal{P} and for all $p \in \mathcal{P}$ the following inclusion holds:*

$$T_p(\mathcal{P}) \subseteq \Delta(\mathcal{P}). \quad (9)$$

Proof. To show that $h_{\mathbf{A}}$ is a smooth embedding we need to prove (i) that it is smooth, which follows from linearity, (ii) that it is a topological embedding, which follows from the assumed injectivity and Prop.10 and (iii) that it has an injective derivative everywhere.

Due to the linearity of $h_{\mathbf{A}}$ on \mathbb{R}^{d^2} we have $h_* = h_{\mathbf{A}}$ but we have to be careful with the domains in order to argue that the injectivity of $h_{\mathbf{A}}$ (as a mapping from \mathcal{P}) implies the injectivity of h_* (as a set of mappings from $T_p(\mathcal{P})$ for any $p \in \mathcal{P}$). By assumption, for any $p \in \mathcal{P}$ and $X \in T_p(\mathcal{P})$ we have $X \in \Delta(\mathcal{P})$. Then indeed $h_*(X) = 0$ together with injectivity of $h_{\mathbf{A}}$ implies $X = 0$ since $h_*(X) = \lambda(h(M_1) - h(M_2))$ is zero only if $M_1 = M_2$. \square

We will now prove the inclusion (9) for the subset of states in $\mathcal{S}(\mathbb{C}^d)$ which are proportional to a projection of rank $r \leq d$, i.e., states which are maximally mixed within a subspace of dimension r . This set forms a smooth manifold of real dimension $2r(d-r)$ which is isomorphic to the *complex Grassmannian* manifold $G(r, d-r)$ [28]. $G(1, d-1)$ is the set of pure states.

Lemma 1. *The inclusion $T_p(\mathcal{P}) \subseteq \Delta(\mathcal{P})$ holds for all $p \in \mathcal{P}$ if \mathcal{P} is the complex Grassmannian manifold $G(r, d-r)$ understood as the submanifold in the space of $d \times d$ selfadjoint matrices which consists of all orthogonal projections of rank r .*

Proof. Let us first identify the tangent space at an arbitrary point $P \in \mathcal{P}$ which is now a selfadjoint projection with $\text{tr}[P] = r$. Considering a curve within \mathcal{P} through P given by the unitary orbit $c(t) := e^{iHt} P e^{-iHt}$ for some selfadjoint matrix H and $t \in \mathbb{R}$. The derivative

$$\partial_t c(t) \big|_{t=0} = i[H, P]$$

is an element of $T_p(\mathcal{P})$ and in fact, such derivatives span the entire tangent space

$$T_p(\mathcal{P}) = \{X = X^\dagger | X = i[H, P] \text{ for some } H = H^\dagger\}. \quad (10)$$

In order to see this we have to show that they span a vector space which has the same dimension as the manifold (for which $D(\mathcal{P}) = 2r(d - r)$). To this end, note that there is a one-to-one relation between commutators and block off-diagonal matrices in the sense that we can always write

$$i[H, P] = \begin{pmatrix} 0 & C \\ C^\dagger & 0 \end{pmatrix}, \quad C \in \mathbb{C}^{r \times (d-r)}, \quad (11)$$

in the basis where $P = \mathbb{1} \oplus 0$. So the dimensions match, which verifies Eq.(10).

In a suitable basis any element $X \in T_P(\mathcal{P})$ is such that

$$X = \left[\bigoplus_{i=1}^r \begin{pmatrix} 0 & c_i \\ c_i & 0 \end{pmatrix} \right] \oplus 0_{d-2r}, \quad c_i \geq 0, \quad (12)$$

since Eq.(11) allows us to work with the singular values $\{c_i\}$ of C by transforming $X \mapsto (U \oplus V)X(U \oplus V)^\dagger$ with appropriate unitaries U and V . Setting $\lambda := \max_i c_i$ equal to the operator norm of X we complete the proof if we show that every 2×2 matrix of the form $c\sigma_x$ with $c \in [0, 1]$ is a difference of two projections. This can be seen to be true by taking the difference of two pure qubit states whose Bloch vectors are parameterized by $(c, \pm\sqrt{1-c^2}, 0)$. \square

A special case is the $2d - 2$ dimensional manifold of pure states in \mathbb{C}^d which can be identified with the *complex projective space* \mathbb{CP}^{d-1} . The map from \mathbb{CP}^{d-1} to selfadjoint rank-one projections is itself a smooth embedding. This together with Proposition 11 and the above Lemma implies that non-embedding results for \mathbb{CP}^{d-1} provide lower bounds on \mathfrak{m} . The probably best non-embedding result in this case is by Mayer [26] from which we now obtain the following:

Corollary 4 (Pure state informational completeness – lower bound). *Informational completeness for the set of pure states in \mathbb{C}^d requires*

$$\mathfrak{m} > \begin{cases} 2D - 2\alpha & \forall d > 1, \\ 2D - 2\alpha + 2 & d \text{ odd, and } \alpha = 3 \bmod 4 \\ 2D - 2\alpha + 1 & d \text{ odd, and } \alpha = 2 \bmod 4, \end{cases} \quad (13)$$

where α denotes the number of 1's in the binary expansion of $d - 1$ and $D = 2d - 2$ is the real dimension of the manifold.

If we now combine this non-embedding result with the upper bounds discussed in Subsection 3.2, we obtain a fairly comprehensive picture on the minimal number of measurements required to identify an element of the set of pure states. For dimensions $d \leq 7$ the upper and lower bounds agree. The minimal number for the dimensions 2 – 7 are summarized in the table below.

d	$4d - 5$	$\mathfrak{m}[\mathcal{P}_1]$
2	3	3
3	7	7
4	11	9
5	15	15
6	19	17
7	23	22

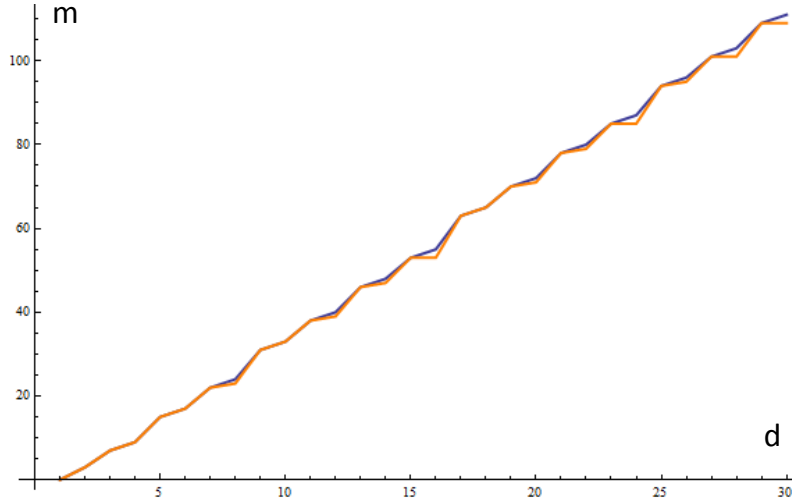


FIGURE 3. Upper and lower bound for the minimal number \mathbf{m} for which informational completeness for pure states in \mathbb{C}^d can be achieved. Note that the bounds coincide for $d \leq 7$ and that they differ by at most 2 until $d = 30$. The slope of the best affine upper bound is exactly 4.

The first gap between the upper and lower bounds occurs for $d = 8$, in which case we know that $23 \leq \mathbf{m}[\mathcal{P}_1] \leq 24$. For all dimensions $d \geq 8$ we can say that the difference between the upper and lower bounds is at most $\log_2(d)$, and that the minimal number differs from the best affine upper bound $4d - 5$ by at most $2\log_2(d)$. The lower and upper bounds up to $d = 30$ are presented in Fig.3.

Remark 7 (Non-embedding results for Grassmannian manifolds). For the $2r(d - r)$ dimensional complex Grassmannian manifold $G(r, d - r)$, i.e., states which are maximally mixed within a subspace of dimension r , again Proposition 11 and the subsequent Lemma assure that non-embedding results carry over to lower bounds on \mathbf{m} . From [27] we obtain a bound for embeddings $G(r, d - r) \rightarrow \mathbb{R}^{\mathbf{m}}$ in the form of Eq.(13) where D has to be set to $D = 2r(d - r)$ and α has to be replaced by $\sum_{j=1}^r \beta(d - j) - \beta(j - 1)$ where $\beta(n)$ is the number of ones in the binary expansion of n . For $r = 1$ this coincides with the aforementioned bound for pure states. Note that this provides a lower bound for informationally complete measurement schemes w.r.t. all sets which includes such a “Grassmannian manifold”, like the set of density matrices with rank bounded by r .

Acknowledgements: MMW was supported by the EU-STREP projects COQUIT and QUEVADIS and the Alfried Krupp von Bohlen und Halbach-Stiftung. TH was supported by the Emil Aaltonen Foundation, Alfred Kordelin Foundation and Academy of Finland (grant no. 1381359). TH and MMW are both grateful for the hospitality and the inspiring working environment at the Institut Mittag-Leffler in Stockholm, where this work was partly done.

REFERENCES

- [1] S. Weigert. Pauli problem for a spin of arbitrary length: A simple method to determine its wave function. *Phys. Rev. A*, 45:7688–7696, 1992.

- [2] J.-P. Amiet and S. Weigert. Reconstructing a pure state of a spin s through three Stern-Gerlach measurements. *J. Phys. A*, 32:2777–2784, 1999.
- [3] J.-P. Amiet and S. Weigert. Reconstructing the density matrix of a spin s through Stern-Gerlach measurements. II. *J. Phys. A*, 32:L269–L274, 1999.
- [4] J. Finkelstein. Pure-state informationally complete and “really” complete measurements. *Phys. Rev. A*, 70:052107, 2004.
- [5] S. Flammia, A. Silberfarb, and C. Caves. Minimal informationally complete measurements for pure states. *Found. Phys.*, 35:1985–2006, 2005.
- [6] D. Gross, Y.-K. Liu, S.T. Flammia, S. Becker, and J. Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, 2010.
- [7] A.S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland Publishing Co., Amsterdam, 1982.
- [8] P. Busch, M. Grabowski, and P.J. Lahti. *Operational Quantum Physics*. Springer-Verlag, Berlin, 1997. second corrected printing.
- [9] E. Prugovečki. Information-theoretical aspects of quantum measurements. *Int. J. Theor. Phys.*, 16:321–331, 1977.
- [10] P. Busch. Informationally complete sets of physical quantities. *Internat. J. Theoret. Phys.*, 30(9):1217–1227, 1991.
- [11] C.M. Caves, C.A. Fuchs, and R. Schack. Unknown quantum states: the quantum de Finetti representation. *J. Math. Phys.*, 43:4537–4559, 2002.
- [12] T. Cubitt, A. Montanaro, and A. Winter. On the dimension of subspaces with bounded Schmidt rank. *J. Math. Phys.*, 49:022107, 2008.
- [13] S.M. Fallat. Bidiagonal factorizations of totally nonnegative matrices. *Amer. Math. Monthly*, 108:697–712, 2001.
- [14] I.M. James. Some embeddings of projective spaces. *Proc. Cambridge Philos. Soc.*, 55:294–298, 1959.
- [15] R.J. Milgram. Immersing projective spaces. *Ann. of Math. (2)*, 85:473–482, 1967.
- [16] A. Mukherjee. Embedding complex projective spaces in Euclidean space. *Bull. London Math. Soc.*, 13:323–324, 1981.
- [17] R. Mañé. On the dimension of the compact invariant sets of certain nonlinear maps. In *Dynamical systems and turbulence*, volume 898 of *Lecture Notes in Math.*, pages 230–242. Springer, 1981.
- [18] B.R. Hunt and V.Y. Kaloshin. Regularity of embeddings of infinite-dimensional fractal sets into finite-dimensional spaces. *Nonlinearity*, 12:1263–1275, 1999.
- [19] J.C. Robinson. Linear embeddings of finite-dimensional subsets of Banach spaces into Euclidean spaces. *Nonlinearity*, 22:711–728, 2009.
- [20] A. Ben-Artzi, A. Eden, C. Foias, and B. Nicolaenko. Hölder continuity for the inverse of Mañé’s projection. *J. Math. Anal. Appl.*, 178:22–29, 1993.
- [21] H. Whitney. The self-intersections of a smooth n -manifold in $2n$ -space. *Ann. of Math.*, 45:220–246, 1944.
- [22] I.M. James. Euclidean models of projective spaces. *Bull. London Math. Soc.*, 3:257–276, 1971.
- [23] M. Adachi. *Embeddings and immersions*, volume 124 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1993. Translated from the 1984 Japanese original by Kiki Hudson.
- [24] S. Chern. On the multiplication in the characteristic ring of a sphere bundle. *Ann. of Math.*, 49:362–372, 1948.
- [25] M.F. Atiyah and F. Hirzebruch. Quelques théorèmes de non-plongement pour les variétés différentiables. *Bull. Soc. Math. France*, 87:383–396, 1959.
- [26] K.H. Mayer. Elliptische Differentialoperatoren und Ganzzahligkeitssätze für charakteristische Zahlen. *Topology*, 4:295–313, 1965.
- [27] T. Sugawara. Nonimmersion and nonembedding theorems for complex Grassmann manifolds. *Proc. Japan Acad. Ser. A Math. Sci.*, 55:59–64, 1979.
- [28] I. Dimitrić. A note on equivariant embeddings of Grassmannians. *Publ. Inst. Math. (Beograd)*, 59:131–137, 1996.

‡ TURKU CENTRE FOR QUANTUM PHYSICS, DEPARTMENT OF PHYSICS AND ASTRONOMY, UNIVERSITY OF TURKU, FINLAND. EMAIL: TEIKO.HEINOSAARI@UTU.FI.

‡ DEPARTMENT OF INFORMATION ENGINEERING, UNIVERSITY OF PADOVA, VIA GRADENIGO 6/B, 35131 PADOVA, ITALY. EMAIL: MAZZARELLA@DEL.UNIPD.IT.

‡ DEPARTMENT OF MATHEMATICS, TECHNISCHE UNIVERSITÄT MÜNCHEN, 85748 GARCHING, GERMANY. EMAIL: WOLF@MA.TUM.DE.